JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **58**, 572–577 (1977)

Inertial Properties of Eigenvalues, II

R. C. THOMPSON

Institute for Algebra and Combinatorics, University of California, Santa Barbara, California 93106

Submitted by L. Mirsky

Inequalities involving the eigenvalues of conjunctive Hermitian matrices are established, and shown to contain a recent result of Machover, the law of inertia, and the interlacing inequalities.

1. INTRODUCTION

Let A and B be $n \times n$ Hermitian matrices, with eigenvalues $\alpha_1 \geqslant \cdots \geqslant \alpha_n$ and $\beta_1 \geqslant \cdots \geqslant \beta_n$, respectively. Let S be a nonsingular $n \times n$ matrix, with singular values $s_1 \geqslant \cdots \geqslant s_n$. Suppose that $B = S^*AS$ is conjunctive with A, where the asterisk indicates transposition and complex conjugation. The law of inertia asserts that A and B have the same numbers of positive eigenvalues, of negative eigenvalues, and of zero eigenvalues. This law was given a quantitative form by Ostrowski [2], who showed that

$$\begin{aligned} \alpha_i s_n^2 \leqslant \beta_i \leqslant \alpha_i s_1^2, &\qquad \text{if} \qquad \alpha_i \geqslant 0, \\ \alpha_i s_1^2 \leqslant \beta_i \leqslant \alpha_i s_n^2, &\qquad \text{if} \qquad \alpha_i \leqslant 0. \end{aligned} \tag{1}$$

In a second paper [3], Ostrowski gave an extension to the case of singular S, and indeed, to the case when S is rectangular, say $n \times m$, with B now $m \times m$ and A still $n \times n$.

Recently [4] Thompson showed, in the case that A, B are both $n \times n$, that many inequalities exist comparing the positive (negative) eigenvalues of B with the positive (respectively, negative) eigenvalues of A and the singular values of S, when A and $B = S^*AS$ are conjunctive. In fact, a general prescription was given in [4] for deriving such inequalities from known inequalities for singular values of matrix products. However, the case of rectangular S and A, B of different dimensions was not considered in [4], this case not seeming to be particularly significant when [4] was written.

However, an interesting paper involving the rectangular case just mentioned was recently published by Machover [1]. Consider the Hermitian form x^*Ax, where $x = (x_1, \ldots, x_n)^T$ is a column n-tuple, and subject the components of x

ISSN 0022-247X

to a set of linear constraints, so that all of $x_1, \ldots, x_n$ are expressed in terms of a subset, say in terms of $y_1, \ldots, y_m$ (certain of $x_1, \ldots, x_n$). Then a new Hermitian form y^*By is obtained, where $y = (y_1, \ldots, y_m)^T$ and where B is $m \times m$ but generally not a principal submatrix of A. If $\alpha_1 \geqslant \cdots \geqslant \alpha_n$ and $\beta_1 \geqslant \cdots \geqslant \beta_m$ are the eigenvalues of A and B, respectively, Machover proved that $\beta_j \leqslant \alpha_j$ provided $\alpha_j \leqslant 0$, and also gave an example showing that we may have $\beta_j > \alpha_j$ if $\alpha_j > 0$.

Although the form of Machover's result suggests a connection with the law of inertia, his inequality is not a direct consequence of Ostrowski's inequalities (1) (although it can be obtained indirectly from (1), see below), nor does it follow from Ostrowski's results [3] in the rectangular case. As Thompson's results [4] are for the nonrectangular case, they do not immediately give Machover's result either.

The chief objective of this paper is to prove a previously unnoticed set of four inequalities generalizing the four inequalities (1). From these inequalities Machover's result will be evident. Though simple to state and easy to prove, our inequalities have a surprising amount of content, since they contain as special cases both (i) the law of inertia and (ii) the interlacing inequalities for the eigenvalues of principal submatrices of Hermitian matrices, as well as (iii) Machover's inequality and (iv) Ostrowski's generalization of (1). Our inequalities are modeled on the inequalities of Weyl for the eigenvalues of sums of Hermitian matrices; as such they have a not unexpected form, their significance being the just mentioned amount of information they contain.

2. The Main Result

Let A be an $n \times n$ Hermitian matrix with eigenvalues $\alpha_1 \geqslant \cdots \geqslant \alpha_n$, and B an $m \times m$ Hermitian matrix with eigenvalues $\beta_1 \geqslant \cdots \geqslant \beta_m$. Let S be $n \times m$, with $s_1 \geqslant \cdots \geqslant s_m$ the singular values of S, i.e., the eigenvalues of the positive semidefinite matrix $(S^*S)^{1/2}$. Suppose that $B = S^*AS$. We have:

THEOREM 1. *If* $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, $i + j - 1 \leqslant m$, *then*

$$\beta_{i+j-1} \leqslant s_i^2 \alpha_j, \qquad \text{when} \qquad \alpha_j \geqslant 0, \tag{2}$$

$$\beta_{i+j-1} \leqslant s_{m+1-i}^2 \alpha_j, \qquad \text{when} \qquad \alpha_j \leqslant 0. \tag{3}$$

THEOREM 2. *If* $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant n$, $i + j > n$, *then*

$$\beta_{i+j-n} \geqslant s_i^2 \alpha_j \qquad \text{when} \qquad \alpha_j \geqslant 0, \tag{4}$$

$$\beta_{i+j-n} \geqslant s_{m+1-i}^2 \alpha_j \qquad \text{when} \qquad \alpha_j \leqslant 0. \tag{5}$$

Proof of (2). Let $e_1, \ldots, e_n$ be orthonormal column eigenvectors of A corresponding, respectively, to $\alpha_1, \ldots, \alpha_n$, let $f_1, \ldots, f_m$ be orthonormal column eigenvectors of B corresponding, respectively, to $\beta_1, \ldots, \beta_m$, and let $g_1, \ldots, g_m$ be orthonormal column eigenvectors of S^*S corresponding, respectively, to $s_1^2, \ldots, s_m^2$.

We require the following lemma.

LEMMA. *Let $S: \mathscr{V} \to \mathscr{W}$ be a linear transformation from vector space $\mathscr{V}$ to vector space $\mathscr{W}$, and let $\mathscr{F}$ be a subspace of $\mathscr{W}$. Then $\mathscr{V}$ contains a subspace $\mathscr{E}$ with $S(\mathscr{E}) \subseteq \mathscr{F}$ and $\dim \mathscr{E} = \dim \mathscr{F} - \dim \mathscr{W} + \dim \mathscr{V}$, provided this last number is nonnegative.*

Proof. Select a basis $w_1, w_2, \ldots$ for $\mathscr{F} \cap$ range S, and let $v_1, v_2, \ldots$ be vectors in $\mathscr{V}$ such that $Sv_1 = w_1$, $Sv_2 = w_2, \ldots$. Let $\mathscr{K}$ be the kernel of S, and take $\mathscr{E}'$ to be the subspace of $\mathscr{V}$ generated by $v_1, v_2, \ldots$ and $\mathscr{K}$. The dimension of $\mathscr{E}'$ is

$$\begin{aligned}
&\dim \mathscr{F} \cap \text{range } S + \dim \mathscr{K} \\
&\quad = \{\dim \mathscr{F} + \dim \text{range } S - \dim(\mathscr{F} + \text{range } S)\} + \{\dim \mathscr{V} - \dim \text{range } S\} \\
&\quad \geqslant \dim \mathscr{F} - \dim \mathscr{W} + \dim \mathscr{V}.
\end{aligned}$$

Take $\mathscr{E}$ to be any subspace of $\mathscr{E}'$ of dimension $\dim \mathscr{F} - \dim \mathscr{W} + \dim \mathscr{V}$.

Now take $\mathscr{V}$ to be the space of column m-tuples, $\mathscr{W}$ to be the space of column n-tuples; and let $\mathscr{F} = \langle e_j, \ldots, e_n \rangle$ be the subspace of $\mathscr{W}$ spanned by $e_j, \ldots, e_n$. By the lemma, a space $\mathscr{E}$ in $\mathscr{V}$ exists such that $S(\mathscr{E}) \subseteq \langle e_j, \ldots, e_n \rangle$ and $\dim \mathscr{E} = n - j + 1 - \dim \mathscr{W} + \dim \mathscr{V} = m - j + 1$. A simple dimensionality calculation now shows that a nonzero vector x exists in

$$\mathscr{E} \cap \langle g_i, \ldots, g_m \rangle \cap \langle f_1, \ldots, f_{i+j-1} \rangle.$$

We may take x to be a unit vector. For this x we have $y = Sx \in \langle e_j, \ldots, e_n \rangle$, and we also have

$$\begin{aligned}
\beta_{i+j-1} \leqslant x^*Bx &= x^*S^*ASx \\
= y^*Ay &\leqslant \alpha_j y^*y \\
= \alpha_j x^*S^*Sx &\leqslant \alpha_j s_i^2,
\end{aligned}$$

the last step following because $\alpha_j \geqslant 0$.

Proof of (3). This proof is precisely the same as that just given, except that we use $\langle g_1, \ldots, g_{m+1-i} \rangle$ in place of $\langle g_i, \ldots, g_m \rangle$.

Proofs of (4) *and* (5). Apply (2) and (3) to $-B = S^*(-A)S$, then replace i and j with $m + 1 - i$ and $n + 1 - j$, respectively.

3. The Theorem of Machover

If we subject the components of the column n-tuple x to a set of linear constraints so that all components of x are expressible in terms of certain components, say the first m components, then

$$x = \begin{bmatrix} I_m \\ M \end{bmatrix} y$$

where $x = [x_1, \ldots, x_n]^T$, $y = [y_1, \ldots, y_m]^T$, I_m is the m-square identity matrix, and M is $(n - m) \times m$. The Hermitian form x^*Ax is converted by this substitution to the form y^*By, where $B = S^*AS$, with

$$S = \begin{bmatrix} I_m \\ M \end{bmatrix}. \tag{6}$$

Since $S^*S = I_m + M^*M$ has 1 plus the (nonnegative) eigenvalues of M^*M as its eigenvalues, it follows that each singular value of S is at least 1. Applying (3) with $i = 1$, we obtain $\beta_j \leqslant s_m{}^2\alpha_j \leqslant \alpha_j$, when $\alpha_j \leqslant 0$, since $s_m \geqslant 1$. This is Machover's theorem. However, as Machover observed, it can happen that $\beta_j > a_j$ when $\alpha_j > 0$. Indeed, if $n - m \geqslant m$, take

$$M = \begin{bmatrix} tI_m \\ 0 \end{bmatrix}, \qquad A = I_n .$$

Then $B = (1 + t^2) I_m$ and $\beta_j > \alpha_j$ for $j = 1, \ldots, m$, if $t > 0$. If $n - m < m$, take $M = [tI_{n-m}, 0]$ and $A = \operatorname{diag}(-I_m, I_{n-m})$, with $t > 0$. Then $B = \operatorname{diag}((-1 + t^2) I_{n-m}, -I_{2m-n})$, and here $\beta_j > \alpha_j$ for all j for which $\alpha_j > 0$, namely $1 \leqslant j \leqslant n - m$, if $t^2 > 2$.

4. The Interlacing Inequalities

Let A be an $n \times n$ Hermitian matrix with eigenvalues $\alpha_1 \geqslant \cdots \geqslant \alpha_n$, and B a principal $m \times m$ submatrix of A, say in leading position, with eigenvalues $\beta_1 \geqslant \cdots \geqslant \beta_m$. We wish to deduce the Cauchy inequalities

$$\alpha_{n-m+j} \leqslant \beta_j \leqslant \alpha_j, \qquad j = 1, \ldots, m,$$

from Theorems 1 and 2. Let

$$S = \begin{bmatrix} I_m \\ 0 \end{bmatrix}.$$

Then $B = S^*AS$. The singular values of S are $s_1 = \cdots = s_m = 1$. By taking

$i = 1$ in either (2) or (3), we deduce that $\beta_j \leqslant \alpha_j$. By taking $i = m$ and substituting $j + n - m$ for j in (4) or (5), we obtain $\beta_j \geqslant \alpha_{j+n-m}$. This shows that the Cauchy inequalities are consequences of Theorems 1 and 2.

5. The Law of Inertia

Taking $m = n$, and S nonsingular, from (4), with $i = n$, it follows that $\beta_j > 0$ whenever $\alpha_j > 0$, and from (3), with $i = 1$, that $\alpha_j > 0$ whenever $\beta_j > 0$. Thus A and B have the same numbers of positive eigenvalues. From (2) and (5) it similarly follows that the numbers of negative eigenvalues coincide. The law of inertia is therefore a consequence of Theorems 1 and 2. (This was already observed by Ostrowski in [2].)

6. Derivation of Machover's Theorem from (1)

This proof is a combination of (1) and the Cauchy inequalities. We have $B = S^*AS$ where S is given by (6). Let

$$S_1 = \begin{bmatrix} I_m & 0 \\ M & tI_{n-m} \end{bmatrix}.$$

There exist unitary matrices U, V such that $UMV = D = [d_{ij}]$ has $d_{ij} = 0$ for $i \neq j$ and $d_{ii} \geqslant 0$ for each i. Then

$$\begin{bmatrix} V^* & 0 \\ 0 & U \end{bmatrix} S_1 \begin{bmatrix} V & 0 \\ 0 & U^* \end{bmatrix} = \begin{bmatrix} I_m & 0 \\ D & tI_{n-m} \end{bmatrix},$$

and after a permutation of rows and the same column permutation the last matrix becomes a direct sum of blocks of the types

$$[1], \qquad [t], \qquad \begin{bmatrix} 1 & 0 \\ d & t \end{bmatrix}.$$

For the third type of block the square of the smallest singular value is

$$\begin{aligned} &\tfrac{1}{2}\{t^2 + 1 + d^2 - ((t^2 + 1 + d^2)^2 - 4t^2)^{1/2}\} \\ &\quad = 2/\{1 + (1 + d^2)\,t^{-2} + ((1 + (1 + d^2)\,t^{-2})^2 - 4t^{-2})^{1/2}\}. \end{aligned}$$

Evidently this smaller singular value is $\leqslant 1$, and approaches 1 as $t \to \infty$. This means that the smallest singular value, call it $\sigma(t)$, of S_1 is at most 1, and approaches 1 as $t \to \infty$. By (1), the jth eigenvalue γ_j of $C = S_1{}^*AS_1$ will satisfy $\gamma_j \leqslant \sigma(t)^2 \alpha_j$ if $\alpha_j \leqslant 0$. As B is a principal submatrix of C, by the Cauchy inequalities we have $\beta_j \leqslant \gamma_j$, hence $\beta_j \leqslant \sigma(t)^2 \alpha_j$. Here $\sigma(t)$ approaches 1 as $t \to \infty$, whereas β_j, α_j remain unchanged. Hence $\beta_j \leqslant \alpha_j$, as was to be shown.

7. Ostrowski's Extension of (1) in the Rectangular Case

When S is $n \times m$, the extension of (1) given in [3] is equivalent to $\beta_j \leqslant s_1{}^2\alpha_j$ if $\beta_j > 0$ and $\beta_{m+1-j} \geqslant s_1{}^2\alpha_{n+1-j}$ if $\beta_{m+1-j} < 0$. The first of these follows immediately from (2) with $i = 1$, since (3) plainly cannot hold when $\beta_j > 0$, $s_m \geqslant 0$, $\alpha_j \leqslant 0$, and the second follows from (4) and (5) in a similar way, by setting m for i and $n + 1 - j$ for j.

8. Further Results

Numerous further inequalities along the lines of those given in [4] can be derived by the method of Section 2.

Acknowledgment

The preparation of this paper was supported, in part, by the U. S. Air Force Office of Scientific Research, under Grant 72-2164.

References

1. M. Machover, Constraints on a single quadratic form, *Trans. New York Acad. Sci.* 2, **35** (1973), 417–420.
2. A. M. Ostrowski, A quantitative formulation of Sylvester's law of inertia, *Proc. Nat. Acad. Sci. U.S.A.* **45** (1959), 740–744.
3. A. M. Ostrowski, A quantitative formulation of Sylvester's law of inertia II, *Proc. Nat. Acad. Sci. U.S.A.* **46** (1960), 859–862.
4. R. C. Thompson, Inertial properties of eigenvalues, *J. Math. Anal. Appl.* **41** (1973), 192–198.